Document Type:	Statistical Analysis Plan
Official Title:	A Phase 2b, Randomized, Double-Blind, Saline Placebo- and Active-Controlled, Multicenter Study of HTX-011 via Infiltration for Postoperative Analgesia in Subjects Undergoing Total Knee Arthroplasty
NCT Number:	NCT03015532
Document Date:	05-Jan-2018

HTX-011-209

A PHASE 2B, RANDOMIZED, DOUBLE-BLIND, SALINE PLACEBO- AND ACTIVE-CONTROLLED, MULTICENTER STUDY OF HTX-011 VIA INFILTRATION FOR POSTOPERATIVE ANALGESIA IN SUBJECTS UNDERGOING TOTAL KNEE ARTHROPLASTY

05 January 2018

Statistical Analysis Plan

Version 4.0

Prepared by:

Heron Therapeutics, Inc.
4242 Campus Point Court, Suite 200
San Diego, CA 92121
USA

Upon review of this document, the undersigned approves the statistical analysis plan. The analysis methods are acceptable, and the table, listing, and figure shell production can begin.

Approved by: [redacted] Date: _____

Vice President, Biometrics
Heron Therapeutics, Inc.

Approved by: [redacted] Date: _____

Vice President, Clinical Research
Heron Therapeutics, Inc.

TABLE OF CONTENTS

List of Abbreviations

Abbreviation	Term
AE	Adverse event
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
ANOVA	Analysis of variance
AST	Aspartate aminotransferase
ATC	Anatomic Therapeutic Classification
AUC	Area under the curve
bpm	Beats per minute
BMI	Body mass index
BUN	Blood urea nitrogen
CI	Confidence interval
CRO	Contract Research Organization
CSR	Clinical Study Report
CTM	Clinical trial materials
DBP	Diastolic blood pressure
DM	Data management
DSMB	Data Safety Monitoring Board
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
eDISH	Evaluation of drug-induced serious hepatotoxicity
GGT	Gamma glutamyltransferase
HCl	Hydrochloride
HR	Heart rate
IM	Intramuscular
IRC	Interim Revivew Committee
IRS	Integrated Rank of Silverman
ITT	Intent-to-Treat
IWRS	Interactive web response system
IV	Intravenous(ly)
K-M	Kaplan-Meier
LDH	Lactic dehydrogenase
LLN	Lower limit of normal
LOCF	Last observation carried forward
LSMD	Least-squares mean difference
MedDRA	Medical Dictionary for Regulatory Activities
MME	Morphine milligram equivalency
MPADSS	Modified Postanaesthetic Discharge Scoring System
NRI	Non-responder imputation
NRS	Numeric Rating Scale
NRS-A	Numeric Rating Scale with activity
NRS-R	Numeric Rating Scale at rest
NSAID	Nonsteroidal anti-inflammatory drug
OBAS	Overall benefit of analgesia score
ORAE	Opioid-related adverse event
PACU	Post-Anesthesia Care Unit
PGA	Patient's Global Assessment
PK	Pharmacokinetic(s)

Abbreviation	Term
PO	Administered orally
PR	Per rectum
PRN	As needed
PT	Preferred term
RBC	Red blood cell
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SBP	Systolic blood pressure
SD	Standard deviation
SDA	Study drug administration
SDTM	Study Data Tabulation Model
SDTMIG	Study Data Tabulation Model Implementation Guide
SE	Standard error
SEM	Standard error of the mean
SI	Standard international
SOC	System Organ Class
SPI	Summed pain intensity
TEAE	Treatment-emergent adverse event
TKA	Total knee arthroplasty
ULN	Upper limit of normal
US	United States
WBC	White blood cell
WHODrug	World Health Organization Drug classification dictionary
wWOCF	Windowed worst observation carried forward

1. ADMINISTRATIVE STRUCTURE

1.1. Sponsor and Oversight

This study is being conducted under the sponsorship of Heron Therapeutics, Inc. (Heron).

[redacted]

1.2. Data Quality Assurance

The Clinical Operations, DM, and Biostatistics departments at the CROs will collaborate internally and with the Sponsor to ensure that the data collected and analyzed for this study are of the highest quality possible and meet the data standards set for the study. This will be accomplished in part through programmed edit checks which will be reviewed by the data managers, statisticians, programmers, and other team members on an ongoing basis to evaluate whether any checks need to be added or any existing checks need to be modified. In addition, periodic blinded reviews of listings of accumulating data, assessment of data query trends, and resulting retraining of study site personnel will be performed to further ensure data quality.

2. INTRODUCTION

This Statistical Analysis Plan (SAP) presents a detailed plan of the statistical methods to be used during the reporting and analysis of efficacy and safety data collected in this study. This SAP does not include the planned analysis and reporting of the pharmacokinetic (PK) assessments and the Holter assessments in the study. Planned PK analyses will be presented in a separate PK analysis plan and planned Holter analyses will be presented in a separate Holter analysis plan.

This SAP was prepared based on protocol version 6, issued 21 December 2017 and was prepared prior to data analysis to provide full details of analyses to be presented in the Clinical Study Report (CSR), including a technical and detailed elaboration of the statistical analysis methods presented in the protocol. Revisions can be made to this SAP while the study is ongoing; however, it must be finalized prior to database lock. Any deviations from the analysis plan provided in the SAP will be fully documented in the final CSR. See Appendix 1 for revision history of the SAP.

This SAP should be read in conjunction with the study protocol and the electronic Case Report Forms (eCRFs).

3. OBJECTIVES

The primary objective of the study is:

- To compare the efficacy and duration of analgesia achieved following periarticular infiltration of HTX-011 with that of saline placebo in subjects undergoing unilateral total knee arthroplasty (TKA)

The secondary objectives are:

- To compare the efficacy and duration of analgesia for HTX-011 with that of bupivacaine hydrochloride (HCl) without epinephrine in this study population.
- To evaluate additional efficacy parameters, including opioid load, in this study population.
- To evaluate the efficacy and duration of analgesia for HTX-011 administered using different techniques in this study population.
- To characterize the bupivacaine and meloxicam PK profiles following administration of HTX-011 in this study population.
- To assess the safety and tolerability of HTX-011 in this study population.

4. INVESTIGATIONAL PLAN

4.1. Overall Study Design and Plan

HTX-011-209 is a Phase 2b, randomized, double-blind, saline placebo- and active-controlled, multicenter study in subjects undergoing primary unilateral TKA. This study will evaluate the efficacy and safety profile of HTX-011 on clinical outcomes (eg, level of pain, duration of analgesia, avoidance of opioids, etc.) in subjects treated with a single intraoperative dose of study drug administered via infiltration to the surgical site.

Cohort 1

Approximately 60 subjects will be randomized to 1 of the following 4 treatment groups in a 2:2:1:1 ratio:

- HTX-011 200 mg (6.8 mL) via periarticular instillation into the surgical site (20 subjects)
- HTX-011 200 mg (6.8 mL) via a combination of periarticular injection and instillation into the surgical site (20 subjects)
- Saline placebo (6.8 mL) via periarticular injection into the surgical site (10 subjects)
- Bupivacaine HCl without epinephrine 0.25% (125 mg, 50 mL) via periarticular injection into the surgical site (10 subjects)

After at least 80% of the planned subjects in Cohort 1 have completed their 72-hour postoperative assessments and have their pain intensity and opioid use data entered into the eCRF, an interim analysis will be performed. An internal Interim Review Committee (IRC), composed of 1 Sponsor representative from each of the Clinical Research, Biometrics, Pharmacovigilance, Regulatory, and Pharmaceutical and Translational Sciences functions will review unblinded summary-level data in order to make a determination to proceed with the next cohort and to select the dose of HTX-011 that will be studied in the next cohort. Detailed responsibilities of the IRC will be presented in an IRC charter.

Cohort 2

Cohort 2 was planned as an optional cohort. Following a review of the results from Cohort 1, the IRC recommended initiating enrollment in Cohort 2 with a single dose level of HTX-011 400 mg/12 mg (bupivacaine/meloxicam doses; 13.7 mL) in each HTX-011 treatment group.

Approximately 200 subjects will be randomized to 1 of the following 4 treatment groups in a 1:1:1:1 ratio:

- HTX-011 400 mg/12 mg (13.7 mL) via periarticular instillation into the surgical site (50 subjects)
- HTX-011 400 mg/12 mg (13.7 mL) via periarticular instillation into the surgical site and ropivacaine 0.5% (50 mg, 10 mL) via periarticular injection into the surgical site (posterior capsule) (50 subjects)
- Saline placebo (13.7 mL) via periarticular injection into the surgical site (50 subjects)
- Bupivacaine HCl without epinephrine 0.25% (125 mg, 50 mL) via periarticular injection into the surgical site (50 subjects)

Just before being taken to the operating room prior to the start of surgery, each subject in Cohort 2 will be administered pregabalin 150 mg orally (PO) and acetaminophen 1000 mg intravenously (IV) to reduce initial postoperative pain. There will be no interim analyses on data collected during Cohort 2.

It is anticipated that this trial will be performed at approximately 30 sites in the United States (US).

4.2. Assessments

Efficacy assessments will include the following:

- Pain intensity scores using the Numeric Rating Scale (NRS) at rest (NRS-R)
- Pain intensity scores using the NRS with activity (NRS-A)
- Use of opioid rescue medication
- Patient's Global Assessment (PGA) of pain control
- Assessment of ability to participate in rehabilitation sessions
- Assessment of ability to ambulate

- Assessments of discharge readiness per the Modified Postanaesthetic Discharge Scoring System (MPADSS)
- Subject's satisfaction with postoperative pain control
- Overall benefit of analgesia score (OBAS)

Safety assessments will include the following:

- Adverse event (AE) recording
- Clinical safety laboratory tests (hematology and serum chemistry)
- Physical examinations
- Wound healing assessment
- Vital signs collections
- ECG
- Continuous Holter monitoring
- Motor function assessment (timed 20-meter walk test)

See the latest version of the study protocol for the timing of procedures and assessments.

4.3. Endpoints

4.3.1. Efficacy Endpoints

The primary efficacy endpoint is:

- Mean area under the curve (AUC) of the NRS-R pain intensity scores through 48 hours (AUC_{0-48}).

The key secondary efficacy endpoint is:

- Mean AUC of the NRS-R pain intensity scores through 72 hours (AUC_{0-72})

Other secondary efficacy endpoints are:

- Mean total postoperative opioid consumption (in morphine equivalents) through 24, 48, and 72 hours.

[redacted]

4.3.2. Safety Endpoints

The safety endpoints are:

- Incidence of treatment-emergent AEs (TEAEs), serious TEAEs (SAEs), and opioid-related TEAEs (ORAEs) through Day 28.
- Change from Baseline in clinical laboratory results.
- Change from Baseline in Holter data.
- Change from Baseline in vital signs at each assessed timepoint.
- Wound healing assessment at 72 hours, and on Day 10 and Day 28.
- Proportion of subjects able to complete a timed 20-meter walk test unassisted at 6, 12, 24, 48, and 72 hours, and on Day 10.

5. GENERAL STATISTICAL CONSIDERATIONS

Unless specified otherwise, all statistical analyses will be performed using a two-sided hypothesis test at the 5% level of significance. All p-values will be rounded to 4 decimal places. If a p-value is less than 0.0001, it will be reported as "< 0.0001". If a p-value is greater than 0.9999, it will be reported as "> 0.9999".

Continuous data will be presented using descriptive statistics: number of subjects (n), mean, standard deviation (SD), median, minimum and maximum. Descriptive statistics on efficacy measures will also include the standard error of the mean (SEM). Categorical data will be summarized by the number and percent of subjects. Confidence intervals (CI) will be 95% and two-sided, unless otherwise stated. Data will be displayed in all listings sorted by cohort, treatment group, subject number and visit/study day. When count data are presented, the percent will be suppressed when the count is zero in order to draw attention to the non-zero counts. A row denoted "Missing" will be included in count tabulations where necessary to

account for dropouts and missing values. The denominator for all percentages will be the number of subjects in that treatment group within the population of interest, unless otherwise stated. Non-zero percentages will be rounded to one decimal place, except 100% will be displayed without any decimal places. Additional rounding rules are as follows:

- If the original value has 0 decimal places: mean, median, and CI will have one decimal place and SD and SEM will have 2 decimal places.
- If the original value has 1 decimal place: mean, median, and CI will have 2 decimal places and SD and SEM will have 3 decimal places.
- If the original value has 2 or more decimal places: mean, median, CI, SD, and SEM will all have 3 decimal places.

Minimum and maximum will always have the same decimal places as the original measure, up to a maximum of 3 decimal places. The above rounding rules will not be applied to original measures displayed in listings.

Values that are collected with "<" or ">" signs will be analyzed as the numerical value without the sign in tables and figures. In listings, these data will be reported as collected with the sign.

All efficacy and safety data will be collected electronically. Datasets will be created using the Study Data Tabulation Model (SDTM) v. 1.3 or higher, conforming to the SDTM Implementation Guide (SDTMIG) v. 3.1.3 or higher. Datasets, tables, listings, and figures will be programmed using SAS® v. 9.4 or higher. All efficacy and safety data will be listed via the SDTM datasets and selected efficacy and safety data will be listed via programmed listings.

5.1. Sample Size

Cohort 1: The sample size of up to approximately 60 subjects was selected empirically without formal statistical assumptions.

Cohort 2: The mean (SD) AUC_{0-48} in the saline placebo and HTX-011 400 mg/12 mg groups after adjusting for opioid rescue medication use is expected to be approximately 425 (90) and 365 (90), respectively. Using a 2-sample t-test, 50 subjects per treatment group results in approximately 90% power to detect a statistically significant treatment effect with $\alpha = 0.05$, 2-sided.

5.2. Randomization, Stratification, and Blinding

Subjects will be randomized to receive HTX-011 (with or without ropivacaine), bupivacaine HCl, or saline placebo. The randomization will not be stratified. Any subject who is randomized but withdraws from the study prior to study drug administration will be replaced by the next eligible subject. The replacement subject will be assigned to the same treatment group as the subject who withdrew. Subjects will not be aware of the study drug they are

receiving. The site's pharmacy and surgical staff will not be blinded to the treatment assignments because HTX-011 is a colored and viscous liquid in contrast to bupivacaine HCl and saline placebo, and the volume of study drug to be administered varies by treatment group. Once surgery is completed and the subject is transferred to the Postanesthesia Care Unit (PACU), the Investigator and all site staff involved in the safety and efficacy assessments, as well as the clinical staff at the CRO involved with study conduct and data collection will be blinded to treatment assignments until after database lock. For Cohort 1, the Sponsor's study team will also be blinded to the treatment assignments with the exception of the clinical trial materials (CTM) staff and an unblinded statistician (from Agility Clinical) who will perform the randomization and interim analysis data review, but will otherwise be uninvolved in the conduct of the study. For Cohort 2, the Sponsor's study team will also be blinded to the treatment assignments with the exception of the CTM staff, Clinical Research Specialists (observers in surgery), and an unblinded statistician (from Agility Clinical) who will perform the randomization, but will otherwise be uninvolved in the conduct of the study.

An internal IRC consisting of 1 Sponsor representative from each of the Clinical Research, Biometrics, Pharmacovigilance, Regulatory, and Pharmaceutical and Translational Sciences functions will be unblinded to summary-level data during the interim analyses conducted in Cohort 1. The IRC will operate under a written, detailed IRC charter.

The randomization will be based on a blocked algorithm and will be done centrally via an interactive web response system (IWRS).

A subject's treatment group assignment will not be broken until the end of the study unless emergency medical treatment of that subject depends upon knowledge of the assigned treatment.

The Sponsor retains the right to break a subject's treatment code for SAEs that are unexpected and are suspected to be causally related to an investigational product and that potentially require expedited reporting to regulatory authorities.

5.3. Statistical Hypotheses and Multiple Endpoint Handling

Cohort 1 and Cohort 2 will be analyzed separately with no data pooling across cohorts.

In Cohort 1, each arm of HTX-011 will be tested against each control arm. The primary comparison on the primary endpoint for each HTX-011 group will be against the saline placebo group under the following hypotheses:

$$H_0: \mu_{\text{HTX-011}} = \mu_{\text{Saline Placebo}}$$
$$H_a: \mu_{\text{HTX-011}} \neq \mu_{\text{Saline Placebo}}$$

Comparisons of HTX-011 against bupivacaine HCl will be considered secondary comparisons. As there is no prospective statistical power calculation of Cohort 1, there will

be no adjustments made to any hypothesis test due to multiple comparisons or multiple endpoints.

In Cohort 2, the following treatment comparisons will be performed in a hierarchical order for the primary and key secondary endpoints:

1. Mean AUC_{0-48} of the NRS-R pain intensity scores for HTX-011 400 mg/12 mg (13.7 mL) via periarticular instillation and ropivacaine 50 mg via periarticular injection vs. saline placebo.

2. Mean AUC_{0-48} of the NRS-R pain intensity scores for HTX-011 400 mg/12 mg (13.7 mL) via periarticular instillation vs. saline placebo.

3. Mean AUC_{0-72} of the NRS-R pain intensity scores for HTX-011 400 mg/12 mg (13.7 mL) via periarticular instillation and ropivacaine 50 mg via periarticular injection vs. saline placebo.

4. Mean AUC_{0-72} of the NRS-R pain intensity scores for HTX-011 400 mg/12 mg (13.7 mL) via periarticular instillation vs. saline placebo.

In order to account for multiple treatment comparisons on the primary and key secondary endpoints in Cohort 2, a strict testing hierarchy will be applied to control study-wise alpha level at 0.05. If the first treatment comparison is statistically significant ($p \leq 0.05$), then the second treatment comparison will be tested. If the second treatment comparison is statistically significant, then the third treatment comparison will be tested. Sequential testing will continue in this manner down the hierarchical order until a treatment comparison fails to meet statistical significance, after which all subsequent treatment comparisons will be considered exploratory.

5.4. Analysis Populations

5.4.1. Intent-to-Treat (ITT) Population

The ITT Population will consist of all subjects who are randomized and receive study drug. This analysis population will be used as the primary analysis population for all efficacy endpoints. The randomized treatment assignment will be used for analysis in this population.

5.4.2. Safety Population

The Safety Population will consist of all subjects who receive study drug. This population will be used for all summaries of safety data. The actual treatment received will be used for analysis in this population.

5.5. Other Important Considerations

5.5.1. Definition of Baseline

Baseline data are defined as the last data collected, whether scheduled or unscheduled, prior to the start of study drug administration.

5.5.2. Calculation of Change and Percent Change from Baseline

Change from Baseline to any timepoint t (C_t) is calculated as follows:

$C_t = M_t - M_B$, where:

- M_t is the measurement of interest at timepoint t
- M_B is the measurement of interest at Baseline

Percent change from Baseline to any timepoint (P_t) is calculated as follows:

$P_t = 100*(C_t/M_B)$

5.5.3. Study Day Calculation for Reporting Purposes

The start of HTX-011 administration will be considered as Time 0 whether or not administered with ropivacaine. The following convention will be used to calculate study day for reporting purposes:

- The study day of study drug administration is Study Day 1.
- For measurements that are *on or after* the date of study drug administration:
  - Study Day = date of measurement – date of study drug administration + 1
- For measurements that are *prior* to the date of study drug administration:
  - Study Day = date of measurement – date of study drug administration

For all subjects, the day of study drug administration should be the same day as the day of the surgical procedure.

5.5.4. Visit Windows

Due to the short duration of the study and the primary efficacy analyses occurring during the 3-day postoperative period of subject hospitalization, no programmatically calculated visit windows are defined for this study.

5.5.5. Handling of Missing and Partial Data

The amount of missing data during the 3-day postoperative primary efficacy analysis period is expected to be very low due to the protocol-required 3-day hospitalization of all subjects following surgery. For any data that is missing, the NRS pain intensity scores will be imputed via last observation carried forward (LOCF), in which the most recent postdose nonmissing value is used for a subsequent missing value. If there is no postdose value

available prior to the first missing value, then the median of values from subjects with nonmissing values within the same treatment group at the relevant timepoint will be used. Predose values will not be carried forward to postdose timepoints.

For binary endpoints (those involving proportions of subjects) not involving the NRS-R or NRS-A, any subject with missing data at a timepoint will be considered as not meeting the endpoint at that timepoint. This is known as nonresponder imputation (NRI). Binary endpoints involving the NRS-R or NRS-A (such as proportion of subjects who are pain-free) through 72 hours will be constructed following windowed worst observation carried forward (wWOCF) (see Section 9.1.1 for details). Binary endpoints involving NRS-R or NRS-A on Day 10 or Day 28 (such as proportion of subjects with NRS score < 4 on Day 10) will be constructed using NRI.

A table displaying the number and percentage of subjects with missing NRS-R pain intensity scores at each nominal timepoint will be produced.

For median time in hours to first opioid rescue administration, subjects who complete the 72 hour observation period without receiving an opioid or discontinue from the study prior to 72 hours without receiving an opioid will be censored at the time of completion or discontinuation, whichever is earlier.

All safety results will be summarized using observed cases with no imputation.

For partial dates involving AE start dates and concomitant medication start dates, the algorithms for imputation will vary depending upon the parameter; the details can be found in Appendix 1. No other partial dates will be imputed.

6. SUBJECT DISPOSITION

A summary of disposition of subjects will include the number and percentage of subjects for the following categories: subjects enrolled (signed the Informed Consent Form), subjects who failed screening with reasons for screen failure, subjects randomized, subjects in the ITT Population, subjects in the Safety Population, subjects completing the 72-hour postoperative observation period, subjects completing study (Day 28), subjects withdrawn from study with the primary reason for withdrawal. Only 1 reason for study withdrawal will be recorded for each subject.

7. DEMOGRAPHICS, CHARACTERISTICS, AND MEDICAL HISTORY

7.1. Demographics and Baseline Characteristics

The demographics and baseline characteristics will be presented in tables using descriptive statistics. The demographics consist of age, age category, sex, race, and ethnicity. The baseline characteristics consist of weight, height, body mass index (BMI), nicotine status, and alcohol use. A subject's age in years is calculated using the integer part of the difference in number of days between the date that informed consent is signed and date of birth divided by 365.25, or is recorded directly on the eCRF. The number and percentage of subjects in the following age categories will be presented: 18-44, 45-54, 55-64, 65-74, 75-84, and ≥85 years old.

Demographics and baseline characteristics will be presented for the Safety Population and demographics only will be presented for all subjects enrolled who fail screening.

7.2. Medical History

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA®) version 19.1. Medical history will be summarized for the Safety Population and will display the number and percentage of subjects with a past and/or concomitant disease or past surgeries by body system.

7.3. Protocol Deviations

Deviations and violations from the protocol will be recorded. Protocol deviations will be classified into, but not necessarily limited to, the following categories:

- ICF procedures
- Eligibility criteria
- Prohibited concomitant medication/therapy
- Laboratory assessment
- Study procedure (eg, efficacy ratings, ECG, PE, etc.)
- Safety reporting
- Randomization/blinding
- Study drug dosing/administration
- Visit schedule/windows

Classification of deviations as important protocol violations will be decided on a case-by-case basis without knowledge of the treatment assigned and before database lock. Protocol deviations will be presented in a summary table by protocol deviation category and treatment. Important protocol deviations will also be listed separately from all protocol deviations.

8. TREATMENTS AND MEDICATIONS

8.1. Prior and Concomitant Medications

Prior medications are defined as medications with a stop date occurring before Study Day 1. Concomitant medications are defined as medications that are ongoing on Study Day 1 or with a start date occurring on or after Study Day 1. Medications with start and stop dates which bracket Study Day 1, or for which missing start and/or stop dates make it impossible to determine the prior or concomitant status, will be summarized as concomitant medications.

All medications will be coded with the World Health Organization Drug classification dictionary (WHODrug).

Prior and concomitant medications will be summarized separately by drug class and preferred term (PT). At each level of summarization, a subject is counted once if that subject reports 1 or more medications at that level. Drug class will correspond to the Anatomical Therapeutic Classification (ATC) Level 2 term.

All prior medications and concomitant medications will be summarized for the Safety Population.

8.2. Surgery Procedure

The knee subject to the surgical procedure (left or right) and the duration of surgery will be summarized. Duration will be calculated as completion time minus start time, reported in minutes.

8.3. Study Drug

For all subjects, treatment will consist of a single intraoperative dose of study drug. As such, extent of exposure will be reported in the CSR as the number of subjects by treatment received in the Safety Population. A summary of treatment compliance will not be produced, as by definition it will be 100% for the Safety Population.

9. EFFICACY ANALYSIS

All efficacy analyses will be performed on the ITT Population. Table 1 displays the planned cohorts and treatment groups being studied.

Table 1. Planned Cohorts and Treatment Group Designations

Cohort	Treatment	Sample Size
1	HTX-011 200 mg/6 mg (6.8 mL) Instillation	20
	HTX-011 200 mg/6 mg (6.8 mL) Combination	20
	Saline placebo (6.8 mL) Injection	10
	Bupivacaine HCl 125 mg (50 mL) Injection	10

Cohort	Treatment	Sample Size
2	HTX-011 400 mg/12 mg (13.7 mL) Instillation + Ropivacaine 50 mg Injection	50
	HTX-011 400 mg/12 mg (13.7 mL) Instillation	50
	Saline placebo (13.7 mL) Injection	50
	Bupivacaine HCl 125 mg (50 mL) Injection	50

See Section 5.3 for details on hypothesis tests and multiple hypothesis handling for primary and key secondary endpoints in Cohort 2.

9.1. Primary Efficacy Endpoint

The primary efficacy endpoint is the mean AUC of the NRS-R pain intensity scores through 48 hours ($AUC_{0\text{-}48}$).

9.1.1. Primary Analysis

During the first 48 hours following surgery, the NRS-R is measured at hours 1, 2, 4, 6, 8, 12, 24, 36, and 48 hours. Using the trapezoidal rule and letting P_t = the NRS-R pain intensity score at time t, then:

$$(t - t_{-1})\frac{P_{t_{-1}} + P_t}{2}$$

is the trapezoidal area between times t and t_{-1}. The $AUC_{0\text{-}48}$ is thus calculated as follows:

$$AUC_{0-48} = \int_{1}^{48} f(t)dt \approx \sum_{i=2}^{48}(t_i - t_{i-1})\frac{P_{i-1} + P_i}{2}$$

The primary endpoint of mean $AUC_{0\text{-}48}$ of the NRS-R pain intensity scores will be analyzed using an analysis of variance (ANOVA) model with randomized treatment as the main effect. In Cohort 1, each HTX-011 group will be tested against each of the control groups. In Cohort 2, treatment comparisons will be performed in a hierarchical order as described in Section 5.3. Results will be expressed as mean AUCs, SDs, and least-squares mean differences (LSMD) and standard errors (SE) with associated 95% CIs, and p-values.

To adjust for the duration effect of opioid rescue medication, the windowed worst observation carried forward (wWOCF) method will be implemented as the primary analysis method for endpoints involving NRS pain intensity scores. In this method, pain intensity scores observed during the analgesic window (duration of effect) of any opioid rescue medication will be replaced with the worst (highest) postdose NRS pain intensity score observed prior to the rescue medication window, with the following exception: if the NRS pain intensity score for a windowed observation is higher than the worst pre-window score, then it will **not** be replaced. wWOCF will be performed following LOCF (ie, perform LOCF

first, and then apply wWOCF). See Table 2 in Section 9.2 for predefined analgesic windows for each opioid medication.

The mean $AUC_{0\text{-}48}$ of the NRS-R pain intensity scores using wWOCF will also be plotted with associated SEM in a bar chart.

9.1.2. Sensitivity Analyses

One sensitivity analysis (using LOCF only) will be performed on the primary endpoint: reproducing the primary analysis but without adjusting the NRS-R pain intensity scores for the use of opioid rescue medications.

[redacted]

9.2.1. Analyses

Key Secondary Endpoint: Mean AUC of the NRS-R pain intensity scores through 72 hours ($AUC_{0\text{-}72}$)

During the first 72 hours following surgery, the NRS-R is measured at hours 1, 2, 4, 6, 8, 12, 24, 36, 48, 60 and 72 hours. The $AUC_{0\text{-}72}$ is thus calculated as follows:

$$AUC_{0-72} = \int_{1}^{72} f(t)dt \approx \sum_{i=2}^{72} (t_i - t_{i-1}) \frac{P_{i-1} + P_i}{2}$$

Similar statistical methods as the primary endpoint will be used for the analysis of this endpoint. See Sections 4.3.1 and 4.3.1 for detail.

The mean $AUC_{0\text{-}72}$ of the NRS-R pain intensity scores using wWOCF will also be plotted with associated SEM in a bar chart.

Secondary Endpoint: Mean total postoperative opioid consumption (in morphine equivalents) through 24, 48, and 72 hours

Determination of morphine equivalents

Use of opioid rescue medication will be examined by unique preferred terms. All opioids used will have the morphine milligram equivalency (MME) calculated (Opioid Morphine Equivalent Conversion Factors, Centers for Disease Control and Prevention, Atlanta, GA, May 2014) based on the conversion factor for each preferred term and the route of the opioid.

Protocol-allowed postoperative rescue medications consist of IV morphine and oral (PO) immediate-release oxycodone. For the first 2 hours in the PACU, the subject may receive up to 15 mg IV morphine, as needed. Thereafter, the subject may receive up to 10 mg in a 2-

hour period, as needed. In addition, the subject may also receive up to 10 mg of PO immediate-release oxycodone within any 4-hour period as needed, from after study drug administration through 72 hours. No other analgesic agents, including ketamine and nonsteroidal anti-inflammatory drugs (NSAIDs), are permitted during the 72-hour postoperative observation period.

Table 2 displays the MME along with the analgesic windows of selected opioid rescue medications for wWOCF purposes. Protocol-allowed postoperative rescue medications are checked. Medications that are not protocol-allowed will be logged as protocol violations, but will still be subject to MME conversion for analysis.

Table 2. Analgesic Windows and Intravenous Morphine Milligram Equivalents for Opioid Rescue Medications

Medication	Route	Window (hr)	MME Factor	Protocol Allowed
CODEINE	PO	6	0.05	
HYDROMOPHONE HYDOCHLORIDE	PO	4	1.33	
HDROMORPHONE HYDOCHLORIDE	IV	4	6.67	
FENTANYL	IV	1	50.00	
HYDROCODONE	PO	6	0.40	
MORPHINE	IV	4	1.00	✓
MORPHINE	PO	4	0.33	
MORPHINE	Intramuscular (IM)	4	1.00	
MORPHINE	Per rectum (PR)	4	1.00	
OXYCODONE	IV	4	1.00	
OXYCODONE	IM	4	1.00	
OXYCODONE	PO	6	0.50	✓
TRAMADOL	IV	6	0.06	
TRAMADOL	PO	6	0.04	

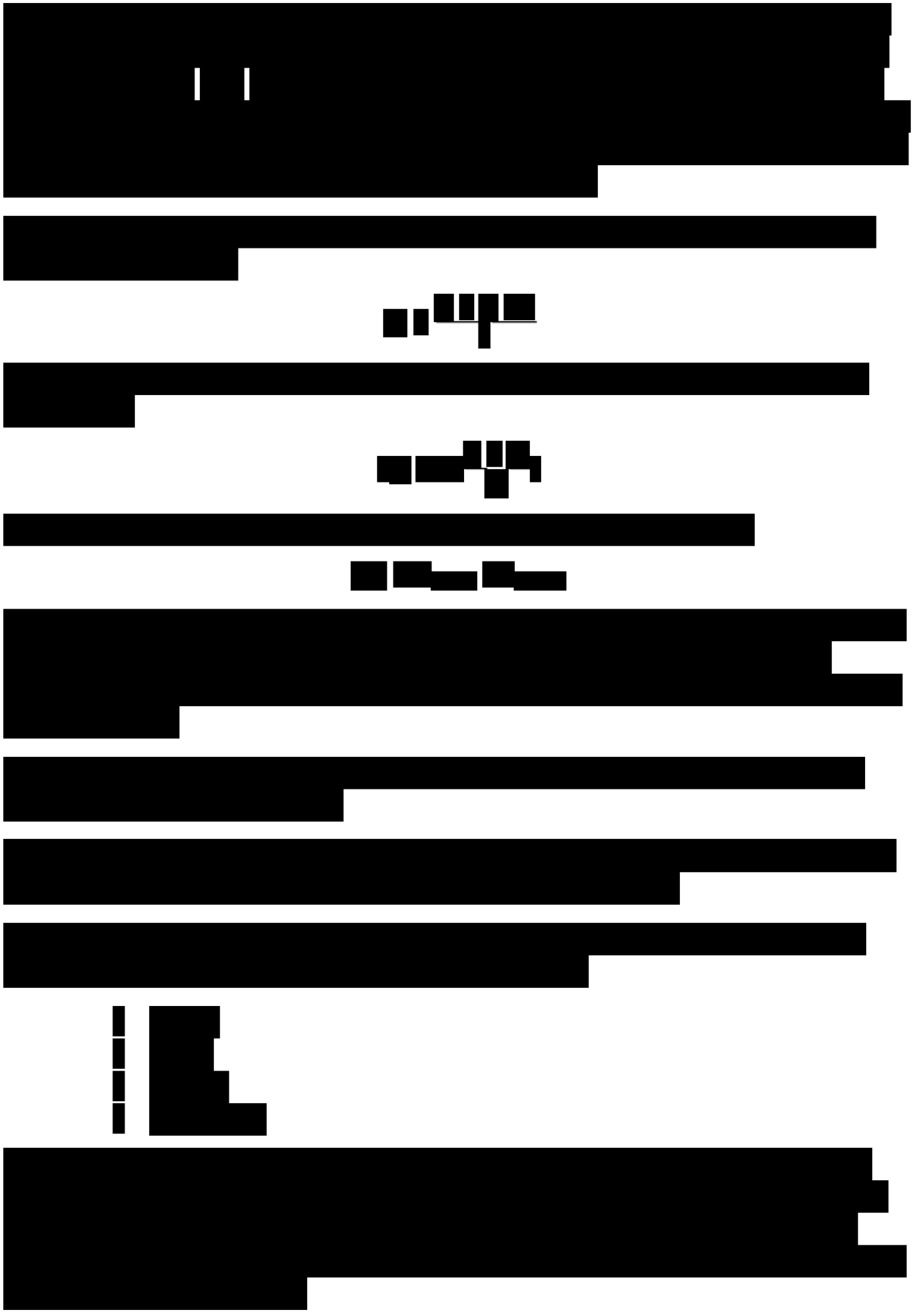
Analysis method

Opioid rescue medication use is collected from hours 0-72. Average daily use and total use will be tabulated using descriptive statistics for each opioid and overall during the following periods: hours 0-24, hours 0-48, and hours 0-72. Subjects who did not use a specific opioid rescue medication during a period of interest will have their dose set to 0 for that period.

The Shapiro-Wilk test will be used to examine the assumption of normality. If this test is statistically significant (ie, $p<0.05$) then the assumption of normality is violated and the total postoperative opioid consumption during each period of interest will be analyzed using Wilcoxon rank sum test. Results will be expressed as median (range) and p-values. However, if assumption of normality holds (ie, Shapiro-Wilk p-value ≥ 0.05), then the total postoperative opioid consumption during each period of interest will be analyzed using an ANOVA model with randomized treatment as the main effect. Results will be expressed as means, SDs, and LSMD and SE with associated 95% CI, and p-values.

The mean total postoperative opioid consumption will also be plotted in a bar chart for each time period with associated SEM.

[redacted]

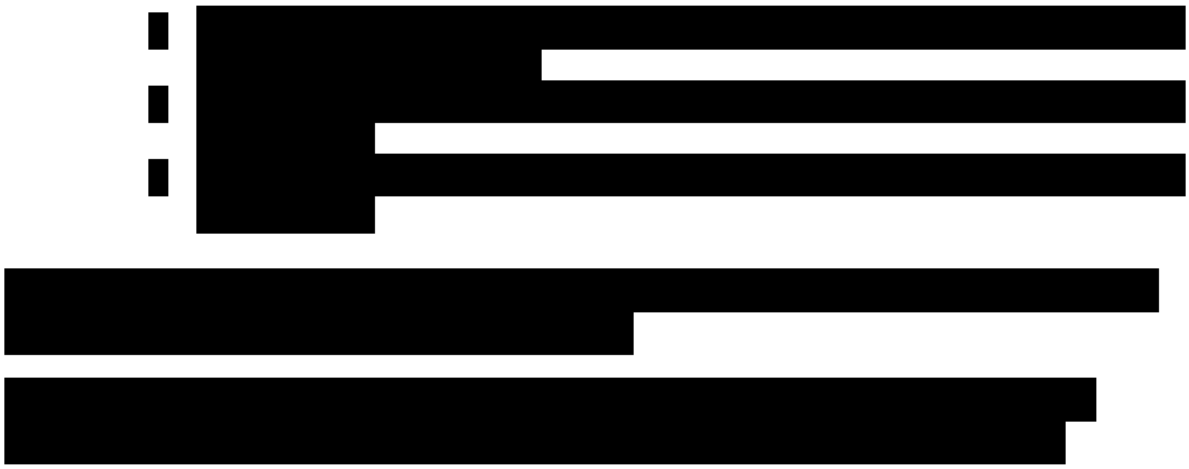


[redacted]

[redacted]

[redacted]

with an NRS-R pain intensity score < 4 at hour 72 who also are < 4 at Day 10



10. SAFETY ANALYSIS

All analyses of safety data will be conducted using the Safety Population. Statistical hypothesis testing will not be performed on any safety results. No imputation of missing safety data will be performed except in the case of partial AE and concomitant medication onset dates (Appendix 1).

10.1. Adverse Events

An AE is defined as any untoward medical occurrence associated with the use of a drug in humans, whether or not considered drug related. A TEAE is any AE which occurs any time during or after study drug administration, or any AE with an onset prior to study drug administration that worsens during or after study drug administration. An AE may be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a study drug, whether or not considered causally associated with the use of the study drug. Any abnormal laboratory value, vital sign result, or ECG finding deemed clinically significant by the Investigator, regardless of causal relationship, must be reported as an AE.

For an event to be a TEAE, it must meet one of the following conditions:

- Begins on Study Day 1, during or after administration of study drug
- Begins after Study Day 1
- Begins before Study Day 1 and worsens in severity during or after the Study Day 1 administration of study drug

AEs with unknown onset dates or unknown end dates will be counted as TEAEs unless the event resolves before Study Day 1.

AEs will be coded using MedDRA version 19.1. Only TEAEs will be presented in AE tables, according to the System Organ Class (SOC), and PT. Any AEs that occur prior to

Study Day 1 will be considered pretreatment AEs and will appear in the AE listing but not in TEAE tables.

10.1.1. Incidence of Treatment Emergent Adverse Events

The incidence of TEAEs table will include only 1 occurrence of a PT per subject. If a subject reports the same PT multiple times, then that PT will only be incremented by 1 since subject counts will be presented. As with the PT, if a subject reports multiple TEAEs within the same SOC, then that SOC will only be incremented by 1 since subject counts will be presented. For tables showing incidence by SOC and PT, SOCs will be sorted by the internationally agreed order and PTs will be sorted within SOC in descending order of incidence in the HTX-011 total column. For tables showing incidence by PT only, the PTs will be sorted in descending order of incidence in the HTX-011 total column.

An overall summary of TEAEs will be presented, and will include the following:

- Number of TEAEs
- Number of subjects with at least 1 TEAE
- Number of subjects with at least 1 possibly related TEAE
- Number of subjects with at least 1 severe TEAE
- Number of subjects with at least 1 TEAE leading to study withdrawal
- Number of subjects with at least 1 opioid-related TEAE (ORAE)
- Number of SAEs
- Number of subjects with at least 1 SAE
- Number of subjects with at least 1 possibly related SAE
- Number of subjects with fatal TEAEs

The incidence of all TEAEs will be presented by SOC and PT and separately by PT only.

10.1.2. Relationship of Adverse Events to Investigational drug

Incidence of possibly related TEAEs to study drug will be presented in a table by SOC and PT. The potential relationships are “Unlikely Related” and “Possibly Related”. TEAEs that are missing relationship will be presented in the summary table as “Possibly Related” but will be presented in the data listing with a missing relationship.

10.1.3. Severity of Adverse Event

Incidence of severe TEAEs will be presented in a table by SOC and PT. TEAEs that are missing severity will be presented in summary tables as “severe” but will be presented in the data listing with a missing severity.

10.1.4. Serious Adverse Events

The seriousness of a TEAE should be assessed by the Investigator independently from the severity of the TEAE. A SAE is an AE occurrence that results in death, is life-threatening,

requires inpatient hospitalization or prolongation of existing inpatient hospitalization, results in persistent or significant disability/incapacity, or is a congenital abnormality/birth defect.

Important medical events that may not be immediately life-threatening or result in death, or require hospitalization may be considered SAEs when, based upon medical judgment, they may jeopardize the subject or may require intervention to prevent one of the outcomes listed above.

Incidence of treatment-emergent SAEs will be presented in a table by SOC and PT. Incidence of possibly related SAEs will also be presented by PT. The incidence of SAE tables will include only 1 occurrence of a PT per subject. If a subject reports the same SAE multiple times, then that PT will only be incremented by 1 since subject counts will be presented. As with the PT, if a subject reports multiple SAEs within the same SOC, then that SOC will only be incremented by 1 since subject counts will be presented. SAEs will also be listed separately.

10.1.5. Adverse Events Leading to Study Withdrawal

All TEAEs reported with "Withdrawal from Study" checked on the eCRF will be presented in a listing.

10.1.6. Opioid-related Adverse Events

Incidence of TEAEs that are potentially opioid-related, regardless of whether a subject actually received an opioid medication, will be presented by PT. Prespecified PTs that may be opioid-related include the following:

- Nausea
- Vomiting
- Constipation
- Pruritus
- Somnolence
- Respiratory depression
- Urinary retention

Incidence of ORAEs will be presented separately as follows:

- Incidence of ORAEs through Day 28
- Incidence of ORAEs through 72 hours
- Incidence of ORAEs through Day 28 in the subset of subjects who received at least 1 opioid rescue medication through 72 hours
- Incidence of ORAEs through 72 hours in the subset of subjects who received at least 1 opioid rescue medication through 72 hours

10.1.7. Death

Any subject deaths during this study will be collected and presented in a listing. The information that is presented will include date of death, days on study, cause of death, and relationship of death to study drug.

10.2. Clinical Laboratory Evaluations

Laboratory assessments will be performed by a central laboratory (hematology and serum chemistry) or locally (pregnancy test and drug screen). All summaries of central laboratory data will be based on the standard international (SI) units provided by the central lab. Associated laboratory parameters such as hepatic profile, renal function, and hematology values will be grouped and presented together.

Summary tables for hematology and chemistry including actual values and change from Baseline values will be presented for clinical laboratory tests with numeric values. These tables will include each visit (Baseline, hour 24 [hematology only], hour 72), highest postdose value, lowest postdose value, and last postdose value.

Laboratory data will also be summarized using shift tables where appropriate. Each subject's hematology and serum chemistry values will be flagged as "low", "normal", or "high" relative to the normal ranges of the central laboratory.

Laboratory data collected at unscheduled visits will be included in listings and will contribute to tables of shifts from Baseline and in tables showing changes from Baseline to highest value, lowest value, and last value. Unscheduled laboratory results will not be windowed for the purposes of assigning a nominal visit.

Listings of laboratory values will include flags for values outside the central laboratory normal ranges that indicate how far out of the normal range an abnormal value is. For example, a value that is ≥ 3 times the upper limit of normal (ULN) but below 4 times the upper limit of normal will have a "3H" flag. Flag multipliers will show values that are 1, 2, 3, 4, 5, and 10 times relative to the ULN if high. Values that are below the lower limit of normal (LLN) will be flagged simply with "L".

Listings of abnormal values for hematology and chemistry will be presented separately in addition to listings of all laboratory values.

10.2.1. Hematology

The following laboratory tests will be included in hematology summary tables: hematocrit, hemoglobin, platelet count, red blood cell (RBC) count, and total and differential white blood cell (WBC) count (basophils, eosinophils, lymphocytes, monocytes, and neutrophils).

10.2.2. Blood Chemistry

The following laboratory tests will be included in the blood chemistry summary tables: alanine aminotransferase (ALT), albumin, alkaline phosphatase (ALP), aspartate aminotransferase (AST), bicarbonate, blood urea nitrogen (BUN), calcium, chloride, creatinine, direct bilirubin, gamma-glutamyltransferase (GGT), glucose, lactate dehydrogenase (LDH), magnesium, phosphorus, potassium, sodium, total bilirubin, total protein, and uric acid.

Associated laboratory parameters such as hepatic profile (ALT, albumin, ALP, AST, direct bilirubin, GGT, total bilirubin), electrolytes (bicarbonate, calcium, chloride, magnesium, phosphorus, potassium, sodium), renal profile (BUN, creatinine), and other (glucose, LDH, total protein, uric acid) will be sorted/grouped together in table and listing presentations.

Evaluation of drug-induced serious hepatotoxicity (eDISH) scatterplots of the highest postdose ALT vs. total bilirubin observed at the same draw as the high ALT value, and of the highest postdose AST vs. total bilirubin observed at the same draw as the high AST value, will be produced.

The incidence of subjects with abnormalities in Liver Function Tests (ALT, AST) will be summarized at each visit for each treatment group for the following categories:

- > 1 x ULN
- ≥ 2 x ULN
- ≥ 3 x ULN
- ≥ 4 x ULN
- ≥ 5 x ULN

10.2.3. Urine Pregnancy Test and Urine Drug Screen

Urine pregnancy test results (women of child-bearing potential) and urine drug screen results will be listed.

10.3. Vital Sign Measurements

Vital signs including systolic blood pressure (SBP), diastolic blood pressure (DBP), heart rate/pulse (HR), and respiration rate will be collected at screening, on Study Day 1 before study drug administration, and at 24, 48, and 72 hours.

Summary tables including actual values and changes from Baseline will be presented for vital signs.

The number and percentage of subjects with clinically relevant abnormalities will be presented using data from any postdose visit (including unscheduled visits). The criteria for clinically relevant abnormalities are shown in Table 3:

Table 3. Clinically Relevant Vital Signs Abnormalities

Vital Sign	Low	High
HR	≤50 bpm and ≥15 bpm decrease from Baseline	≥120 bpm and ≥15 bpm increase from Baseline
SBP	≤90 mmHg and ≥20 mmHg decrease from Baseline	≥160 mmHg and ≥20 mmHg increase from Baseline
DBP	≤50 mmHg and ≥15 mmHg decrease from Baseline	≥100 mmHg and ≥15 mmHg increase from Baseline

Abbreviations: bpm, beats per minute; DBP, diastolic blood pressure; HR, heart rate; SBP, systolic blood pressure.

10.4. Electrocardiogram

12-lead ECG (triplicate) is performed at screening. Screening ECG results will be listed only, and will include each of the 3 triplicate ECGs as well as the mean values from the triplicate assessments.

10.5. Physical Examination

Physical examination is performed at Screening, 72 hours, and early termination. Physical examination results will be listed only.

10.6. Wound Healing Assessment

Wound healing assessment according to the Southampton Wound Scoring System (Bailey, Karran et al. 1992) is performed at 72, and 120 hours, and on Days 10 and 28. A summary of wound healing assessment results will be produced, showing the number and percentage of subjects at each visit by grade with subgrade breakdown. Wound healing assessment results will also be listed.

10.7. Motor Function Assessment

Motor function assessment is performed at Screening, at 6, 12, 24, 48, and 72 hours, and on Day 10. The proportion of subjects able to complete a timed 20-meter walk test unassisted (a successful motor function assessment) will be summarized at each timepoint.

11. INTERIM ANALYSIS

11.1. Interim Analysis

One interim analysis will be performed. An internal IRC composed of 1 Sponsor representative from each of the Clinical Research, Biometrics, Pharmacovigilance, Regulatory, and Pharmaceutical and Translational Sciences functions will review unblinded summary-level data from Cohort 1 in order to make a determination to proceed with the next cohort and to select the dose(s) of HTX-011 that will be studied in the next cohort. The interim analysis will be performed after at least 80% of the planned subjects in Cohort 1 have

had their pain intensity and opioid use data through the 72-hour postoperative assessments entered into the eCRF. Detailed responsibilities of the IRC will be presented in an IRC charter. No adjustments to the type-1 error rate will be made in the efficacy hypothesis testing as a result of this interim analysis.

There will be no interim analyses for Cohort 2.

11.2. Data Safety Monitoring Board

An independent Data Safety Monitoring Board (DSMB) will not be involved with the conduct of this study. The Sponsor will review blinded tables and listings of accumulating data approximately bi-weekly to check enrollment, adherence to follow-up schedule, and ongoing safety results.

12. REFERENCES

Bailey, I. S., S. E. Karran, K. Toyn, P. Brough, C. Ranaboldo and S. J. Karran (1992). "Community surveillance of complications after hernia surgery." BMJ **304**(6825): 469-471.

Chan, I. S. and Z. Zhang (1999). "Test-based exact confidence intervals for the difference of two binomial proportions." Biometrics **55**(4): 1202-1209.

Lehmann, N., G. P. Joshi, D. Dirkmann, M. Weiss, P. Gulur, J. Peters and M. Eikermann (2010). "Development and longitudinal validation of the overall benefit of analgesia score: a simple multi-dimensional quality assessment instrument." Br J Anaesth **105**(4): 511-518.

Silverman, D. G., T. Z. O'Connor and S. J. Brull (1993). "Integrated assessment of pain scores and rescue morphine use during studies of analgesic efficacy." Anesth Analg **77**(1): 168-170.

APPENDIX 1. IMPUTATION OF PARTIAL AND MISSING DATES

Incomplete Dates of Adverse Event start

All AE onset dates must be entered on the eCRF as complete dates. In the rare case that all or part of an AE onset date is missing but an AE resolution date is present and after study drug administration then the AE onset date will be imputed as follows:

Year of onset	Month of onset	Day of onset	Onset date to be imputed as
Missing	Missing	Missing	Date of SDA
year = year of SDA	Missing	Nonmissing	Set month to month of SDA
year = year of SDA	Missing	Missing	Set month and day to those of SDA
year < year of SDA	Missing	Nonmissing	set month to December
year < year of SDA	Missing	Missing	set month and day to December 31
year > year of SDA	Missing	Nonmissing	set month to January
year > year of SDA	Missing	Missing	set month and day to January 1
year = year of SDA	Month = month of first dose	Missing	Set day as day of 1st dose
year = year of SDA	Month < month of first dose	Missing	Set day as last day of onset month
year = year of SDA	Month > month of first dose	Missing	Set day as first day of onset month
year < year of SDA	Nonmissing	Missing	Set day as last day of onset month
year > year of SDA	Nonmissing	Missing	Set day as first day of onset month

SDA = study drug administration.

If AE resolution date is present and prior to study drug administration, there is no need to impute incomplete AE start date, as the AE is not treatment emergent and the event should be in the medical history.

Concomitant Medications

- If year and month are present and day is missing then set day to first day of month for start date, and set day to last day of month for end date
- If year and day are present and month is missing then set month to January for start date, and set month to December for end date
- If year is present and month and day are missing then set month and day to January 1 for start date, and set month and day to December 31 for end date
- Completely missing dates will not be imputed

If start date is completely missing and end date is not prior to study drug administration, then the medication will be classified as concomitant; if the end date is missing, then the medication will be classified as ongoing. Medications for which the start and end dates are completely missing will be classified as concomitant.

APPENDIX 2. DOCUMENT REVISION HISTORY

Version	Date	Notes/Revisions
1.0	02 March 2017	Initial version, based on protocol version 2 (19 January 2017)
2.0	30 March 2017	Second version, based on protocol version 3 (28 March 2017) • Blinded assessor's satisfaction with postoperative pain control removed as an endpoint • Primary analysis of Integrated Rank of Silverman changed from using NRS-R to NRS-A • Modified PGA endpoint incorporated, using a 4-point scale instead of a 5-point scale • Opioid use diary added, enabling endpoints and analyses of opioid use through Day 28 • Opioid-related TEAEs added as an analysis • Primary analysis changed from not adjusting for opioid use to adjusting for opioid use (wWOCF) • Added a series of moderate-severe pain analyses • Added a series of analyses on maintenance of none-mild pain • Other minor changes in study drug volume decimal places and formatting
3.0	18 September 2017	Third version, based on protocol version 5 (28 August 2017) • Removed Cohort 3 • Added HTX-011 dose and sample size chosen for Cohort 2 based on IRC recommendation • Changed primary efficacy endpoint to "Mean area under the curve (AUC) of the NRS-A pain intensity scores through 72 hours (AUC_{0-72})" • Removed summed pain intensity (SPI) related efficacy endpoints • Clarified that Cohort 1 and Cohort 2 would be analyzed separately without pooling across cohorts • Added formal type 1 error control for Cohort 2 (hierarchical testing of comparison for the primary endpoint) • Added formal statistical determination of sample size for Cohort 2 • Clarified Intent-to-Treat Population • Removed Modified Intent-to-Treat Population • Removed interim analysis of data from Cohort 2. • Updated protocol deviation categories
4.0	29 December 2017	Fourth version, based on protocol version 6 (21 December 2017) • Increased the sample size of Cohort 2 from approximately 100 subjects to approximately 200 subjects (50 subjects per treatment group) • Revised the primary endpoint as follows: Numeric Rating Scale of pain intensity scores with activity (NRS-A) was changed to Numeric Rating Scale of pain intensity scores at rest (NRS-R). The time interval was changed from 0-72 hours to 0-48 hours • Added a key secondary endpoint: Mean AUC of the NRS-R pain intensity scores through 72 hours (AUC0-72) • [redacted]

Version	Date	Notes/Revisions
		[redacted] • Revised the analysis plan for Cohort 2 to include a hierarchical order for the primary and key secondary endpoints • Revised the sample size calculation statement in Cohort 2